LEXONOMICA
Let. 7, št. 2, str. 161-188, december 2015

Veliko podatkovje v pravu in ekonomiji: veliki izzivi ali velike težave?

LILJANA SELINŠEK

Povzetek Prispevek obravnava izbrane ekonomske in pravne vidike velikega podatkovja in analizira novejše trende oz. mnenja o potrebnih pristopih k pravni regulaciji tega področja, ki je tesno zvezano s hitrim tehnološkim napredkom. Temeljna teza prispevka je, da razmeroma enovitega in podatkov polnega virtualnega sveta oz. kibernetskega prostora ne bo mogoče uspešno regulirati parcialno po posameznih pravnih področjih, ampak bo tradicionalno delitev pravnih področij treba preseči z novim krovnim (mednarodnim in široko sprejetim) pravnim aktom oz. področjem, ki bo določilo osnovna pravila in pravna načela, po katerih se bodo nato lahko ravnale tudi posamezne pravne discipline. Če želi pravo igrati aktivno in konstruktivno vlogo v spremenjenih družbenih razmerah ter biti dolgoročno upošteven dejavnik oz. razumna protiutež tehnološkemu razvoju, je potreben sistemski pristop na globalni ravni. Izziv je torej velik, a tudi težav ne manjka. Ena bistvenih je čas – vprašanje je, ali bo pravo še zmožno obvladati tehnologijo, ko bo (končno) dojelo, za kaj gre.

KJUČNE BESEDE: • veliko podatkovje • digitalno vesolje • podatkovna pravica • napovedni modeli • varstvo osebnih podatkov • zasebnost • pravo

NASLOV AVTORICE: Dr. Liljana Selinšek, docentka, Inštitut za kriminologijo pri Pravni fakulteti v Ljubljani, Poljanski nasip 2, SI-1000 Ljubljana, Slovenija, e-pošta: liljana.selinsek@ifit.si.

DOI 10.18690/18557147.7.2.61-188(2015), UDK: 346.546.5:347.45, JEL: K12, K21
ISSN 1855-7147 Tiskana izdaja / 1855-7155 Spletna izdaja
Na svetovnem spletu dostopno na http://journals.lexonomica.press.

LEXONOMICA
Vol. 7, No. 2, pp. 161-188, December 2015

LEXONOMICA

Big Data in Law and Economy: Big Challenges or Big Troubles?

LILJANA SELINŠEK

Abstract Article outlines some economic and legal aspects of big data, and provides analysis of recent trends and opinions on necessary approaches to the legal regulation of this area which is tightly connected with fast technological progress. Article sets essential thesis that virtual world or cyberspace - that is relatively uniform and full of data - cannot be constructively regulated partially in each separate legal field. Instead, the traditional division of legal fields or disciplines should be exceeded by a new general (international and commonly accepted) legal act or discipline that will set up basic legal rules and principles to be then adopted in different legal fields. If law aspires to play active and constructive role in new social environment and to be a factor of lasting importance or reasonable counterpoise to technological progress, systemic approach on global level is essential. Therefore, challenges are big, but there are troubles as well. One of the important ones is time – i.e. the question whether the law will be able to master technology when it (finally) realizes what the point is.

KEYWORDS: • big data • digital universe • dataright • predictive policing • personal data protection • privacy • law

CORRESPONDENCE ADDRESS: Liljana Selinšek, Ph.D., Assistant Professor, Institute for Criminology at Law Faculty in Ljubljana, Poljanski nasip 2, SI-1000 Ljubljana, Slovenia, e-mail: liljana.selinsek@ifit.si.

DOI 10.18690/18557147.7.2.161-188(2015), UDK: 346.546.5:347.45, JEL: K12, K21
ISSN 1855-7147 Print / 1855-7155 On-line
Available on-line at http://journals.lexonomica.press.

1 Uvod

Sodobna informacijsko-komunikacijska tehnologija, ki je v zadnjih dveh desetletjih korenito spremenila naš način življenja, ustvarja zanimiv pojav (včasih načrtno, včasih pa zgolj kot stranski produkt vsakdanjih, tudi povsem rutinskih aktivnosti) – ogromne količine podatkov v digitalni obliki, ki so sčasoma dobile lastno poimenovanje: veliko podatkovje oz. *big data*. Temeljne značilnosti družbene realnosti, ki generira veliko podatkovje, so predvsem naslednje (Murphy, Barton, 2014: 8):

- razvoj tehnologije, ki omogoča masivno zbiranje podatkov in rudarjenje po podatkih ob zmernih stroških (za razmah velikega podatkovja je v veliki meri zaslužno vztrajno padanje cen dostopnosti do interneta in višanje zmogljivosti računalniške tehnologije ob hkratnem padanju cen, kar vse omogoča dostopnost do informacijsko-komunikacijske tehnologije čedalje širšemu krogu ljudi, več odjemalcev in uporabnikov pa pomeni več podatkov. Tudi orodja in tehnologija za obdelavo, shranjevanje in analizo podatkov so cenovno čedalje bolj dostopna[1]),
- pojav t. i. interneta stvari (internet of things) oz. povezanih naprav, ki ujamejo, shranijo in prenašajo podatke v realnem času (pametni mobilni telefoni, povezane računalniške naprave, najrazličnejši senzorji, ki merijo vse mogoče, od prometa in vremena do porabe energije). Internet stvari ustvarja ogromne nove vire podatkov, ki sami zase nimajo nujno posebne vrednosti, vendar lahko ustrezno kombinirani in analizirani razkrijejo vzorce obnašanja in druge značilnosti, ki imajo veliko tržno ali družbeno vrednost,
- pojav družbenih omrežij in podobnih virov informacij o posameznikovih zanimanjih, navadah in socialni mreži in
- izjemen razmah internetnih aplikacij in programja za podatkovno analitiko.

Danes je v bistvu vsaka informacija sposobna za digitalizacijo in hrambo, vključno z internetnimi iskanji, spletnimi nakupi, objavami na Facebooku, telefonskimi klici, uporabo cestnine, finančnimi transakcijami in vsemi besedami v knjigah. Po podatkih IBM je bilo 90 % vseh podatkov na svetu ustvarjenih v preteklih dveh letih. Po podatkih Googla ljudje danes v dveh dneh ustvarimo toliko informacij oz. podatkov, kot smo jih prej od začetka civilizacije pa do leta 2003. Nekateri so mnenja, da bo velikost podatkov narastla do te mere, da njihovega obsega ne bo mogoče več številčno opredeliti (povzeto po Joh, 2014: 38), kar je realen scenarij, saj razmaha podatkov ni več mogoče ustaviti. Posamezniki ustvarjajo in bodo ustvarjali čedalje več podatkov s pametnimi mobilnimi telefoni, koriščenjem storitev v računalniškem oblaku, uporabo družbenih omrežij in drugih novih (danes morda še neznanih) tehnologij, podjetja in tudi države pa bodo še naprej zbirale te podatke in iskale nove načine njihove uporabe. Poceni in majhni, a sofisticirani senzorji in sledilne naprave, ki merijo in zbirajo podatke, niso vgrajene le v številne proizvode in naprave, ampak lahko

obvladujejo cela mesta, ki na ta način postanejo »pametna«. Prototip za takšno pametno mesto (smart city) je denimo Santander v Španiji. Po tem mestu je nameščenih več kot 12.000 senzorjev, ki merijo vse - od onesnaženosti zraka do števila prostih parkirnih mest. Zbrani podatki se nalagajo v »podatkovnih skednjih«, ki hranijo vztrajno naraščajoče število informacij o mestu in njegovih prebivalcih in obiskovalcih (Joh, 2014: 39–40), na podlagi katerih je mogoče analizirati in tudi napovedovati najrazličnejše vidike življenja v tem mestu.

Kljub izjemnemu obsegu velikega podatkovja, ki iz minute v minuto narašča, pa je mogoče ugotoviti, da ta fenomen še nima ustreznega mesta v pravu, pa tudi ekonomsko še ni eksplodiral tako, kot je bilo pričakovano. Delno sta ti zadevi povezani, saj se ekonomski in pravni vidiki velikega podatkovja prepletejo pri vprašanju, ali pravo morda ne zavira gospodarskega razvoja s pretiranimi ali nerazumnimi omejitvami. Dokončnih odgovorov na to in tudi številna druga vprašanja, ki jih poraja veliko podatkovje, (še) ni. Če je z ekonomskega zornega kota največji izziv velikega podatkovja v tem, kako podatke izkoristiti za posel, je s pravnega vidika vsaj zaenkrat najpomembnejše vprašanje, kako v dobi velikega podatkovja zaščititi zasebnost in osebne podatke posameznikov oz. ali je zasebnost v digitalni dobi sploh še realen koncept. Dodatne pomembne izzive za pravni okvir sproža vpeljava metodologij, ki temeljijo na velikem podatkovju, v policijsko in tožilsko delo, predvsem z zvezi z mejami oz. omejitvami glede uporabe podatkov za utemeljevanje policijskega nadzora, preiskav in posegov (Joh, 2014: 41). Precej izzivov (in težav) na tem področju je posledica opustitve vsakršne regulacije interneta, ki je eden pomembnih generatorjev spreminjanja družbenega okolja na nacionalni in globalni ravni. Pri tem je bistveno, da interneta, ki je pomemben vir velikega podatkovja in dostopnosti do podatkov, zaradi njegove narave preprosto ni mogoče regulirati na nacionalni ravni oz. z nacionalnim pravom, ampak bi bila potrebna intervencija na področju mednarodnega prava (sorodno kot npr. pri pomorskem pravu).[2]

Povečano zbiranje podatkov in njihova uporaba gotovo imata velik (ekonomski) potencial za izboljšanje učinkovitosti in dodano vrednost na različnih področjih, vendar pa je uresničitev tega potenciala pomembno odvisna od prava, ki mora (po)iskati ustrezno ravnovesje med pravicami potrošnikov, gospodarskimi interesi in raziskovalnimi težnjami (Murphy, Barton, 2014: 15–16, tudi Mattioli, 2014: 540), pa tudi uravnotežiti pravice in potrebe posameznikov na eni in države oz. družbe na drugi strani. Pravni teksti, ki obravnavajo različne vidike velikega podatkovja, so sicer čedalje pogostejši, a večinoma brez pravih (tj. inovativnih in hkrati realnih) idej, kako pravno tradicijo (pa naj si bo anglo-saksonskega ali kontinentalnega tipa) združiti s poplavo podatkov, ki je posledica izjemnega tehnološkega napredka zadnjih nekaj desetletij.

Ta prispevek obravnava izbrane ekonomske in pravne vidike velikega podatkovja in analizira novejše trende oz. mnenja o potrebnih pristopih k pravni regulaciji tega področja. Temeljna teza prispevka je, da razmeroma enovitega virtualnega

sveta oz. kibernetskega prostora, ki prekipeva od podatkov, ne bo mogoče uspešno regulirati parcialno s pravom, razcepljenim na posamezna področja, ampak bo tradicionalno delitev pravnih področij treba preseči z novim krovnim (mednarodnim in široko sprejetim) pravnim aktom (ali področjem), ki bo določilo osnovna pravila in pravna načela, po katerih se bodo nato lahko ravnale tudi posamezne pravne discipline. Če želi pravo igrati aktivno in konstruktivno vlogo v spremenjenih družbenih razmerah ter biti dolgoročno upošteven dejavnik oz. razumna protiutež tehnološkemu razvoju, je torej potreben sistemski pristop na globalni ravni.

2 Fenomen velikega podatkovja

Čeprav ni poenotene definicije velikega podatkovja, se večina avtorjev strinja, da se ta izraz nanaša na aplikacijo umetne inteligence na velik obseg digitaliziranih podatkov (Joh, 2014: 38) oz. da s praktičnega zornega kota izraz veliko podatkovje 1) označuje tehnologijo, ki maksimizira računalniško moč in algoritmično točnost, 2) opisuje vrste analiz, ki temeljijo na orodjih oz. programih[3] za prečiščevanje in primerjavo podatkov in 3) ustvarja mnenje, da velike podatkovne baze zagotavljajo bolj resnične, objektivne in točne rezultate (Crawford, Schultz, 2014: 96). Pri tem je treba poudariti, da v primerjavi z »navadnimi« podatki veliko podatkovje ni le večje po obsegu, ampak tudi drugačno po naravi, pri čemer ga definirajo predvsem naslednje temeljne značilnosti (Dai in ostali, 2012: 93):

- nestrukturiranost (podatki prihajajo iz veliko virov in v zelo različnih oblikah),
- velik obseg in hitra rast (nestrukturirani podatki rastejo 10-15 krat hitreje od strukturiranih in bodo kmalu predstavljali 90 % vseh podatkov. Veliko podatkovje je torej obsežno, njegov obseg pa eksponentno raste, in sicer za približno 60 % letno),
- fleksibilen okvir za upravljanje s podatki in analitiko, temelječ na razpršenem programskem delovanju v grozdih neodvisnih strežnikov,
- napovedna analitika, temelječa na kompleksnem strojnem učenju oz. umetni inteligenci, statističnih modelih, analizah grafov ipd., ki z obdelavo ogromne količine podatkov omogoča identifikacijo bodočih trendov in vzorcev (skoraj) v realnem času, saj podatki kontinuirano pritekajo v sistem.

Izraz veliko podatkovje torej v bistvu označuje novo metodo empiričnega raziskovanja (Mattioli, 2014: 539), ki vključuje črpanje podatkov iz potencialno številnih in različnih podatkovnih baz, ki prvotno niso bile nujno namenjene kombiniranju. Tehnologija, ki omogoča zbiranje in obdelavo ogromne količine podatkov, je zelo raznolika in se nenehno razvija, temelji pa na tem, da se z analitičnimi orodji in procesi preveri, ali med podatki obstajajo pomenske korelacije in povezave. Dodana vrednost nastane, kadar analiza pokaže nek nov, uporaben vpogled (Murphy, Barton, 2014: 8). Veriga oz. življenjski cikel

podatkov je okvirno naslednji (rezultati analize lahko ustvarijo dodatne podatke in odzive v podatkovni verigi, ki lahko sprožijo nov življenjski cikel podatkov) (OECD, 2013: 7):

Izbris kot zadnja faza je označen ločeno, saj je njegov status člena v podatkovni verigi še odprt. Izbris podatkov je pomemben s pravnega vidika (predvsem z zornega kota varstva osebnih podatkov), vendar pa ne sovpada s trenutno veljavnimi tehnološkimi (pa tudi ekonomskimi) koncepti velikega podatkovja, ki temeljijo na tem, da se podatki zbirajo in hranijo čim dlje oz. po možnosti trajno in »na zalogo«, brez vnaprej definiranega namena uporabe in z neomejenimi možnostmi rudarjenja po njih. Ideja velikega podatkovja torej temelji na tem, da se najprej zberejo ogromne količine podatkov, nato pa konstantno iščejo možnosti za kombinacijo, obdelavo in analize. Koliko je posamezen podatek oz. informacija dejansko vredna, namreč postane znano šele, ko je ta podatek mogoče povezati z drugim podatkom, tudi takšnim, ki nastane šele v prihodnosti.

(Trenutno) osrednja infrastruktura, ki omogoča razmah velikega podatkovja, je računalništvo v oblaku (cloud computing), ki tvori pomemben del t. i. digitalnega vesolja (digital universe). To vesolje se izjemno hitro širi, in sicer vse od začetka uporabe interneta naraste v povprečju dvakrat vsako leto. Veliko večino tega vesolja ljudje ustvarimo (in porabimo) z uporabo interneta (vključno z udejstvovanjem na družbenih omrežjih), gledanjem digitalne televizije ipd., 80 % vseh podatkov, ki tvorijo digitalno vesolje, pa obvladujejo spletna podjetja Google, Facebook, Twiter ipd. (Donova, Finn, Wadhva, 2014: 14). Med vodilnimi zbiratelji podatkov so danes pametni mobilni telefoni, ki s kombinacijo geo-lokacijskih podatkov in dostopa do interneta omogočajo oz. podpirajo številne nove storitve in aplikacije, ki so povezane s prometom, okoljem, zdravjem ipd. Mnoge od teh storitev in aplikacij za delovanje potrebujejo osebne podatke ali pa te podatke zbirajo (OECD, 2013: 8).

Veliko podatkovje spreminja tudi vzorce delovanja v znanosti. Veliko znanstvenih področjih že temelji na ogromnem številu podatkov in razvoju računalniške tehnologije. Denimo meteorologija, astronomija, astrofizika, bioinformatika in računalniška biologija so v veliki meri naslonjene na podatkovno intenzivna znanstvena odkritja. Na teh področjih se generirajo oz. ustvarjajo velike količine podatkov najrazličnejših vrst, njihova obdelava oz. uporaba za znanstvene simulacije in analize pa je eden od znanstvenih izzivov moderne dobe. V ta namen je v teku ali razpisanih precej t. i. e-znanstvenih projektov (e-science), v okviru katerih se razvijajo tudi zadeve, kakršen je veliki hadronski trkalnik, ki pri pospeševanju delcev ustvari 60 terabajtov podatkov na dan, zmogljivi vesoljski teleskopi, ki delujejo kot digitalne kamere in prav tako ustvarjajo ogromne količine podatkov dnevno itd. (več o tem Chen, Zhang, 2014: 317).

3 Nekaj ekonomskih vidikov velikega podatkovja

3.1 Splošno

Ekonomisti veliko podatkovje oz. tehnologije, ki ga omogočajo, povezujejo z elementi gospodarnosti in jih definirajo denimo kot novo generacijo tehnologij in infrastrukture, ki je zasnovana tako, da gospodarno izvleče vrednost iz zelo velike količine zelo raznolikih podatkov na način, da omogoča njihovo zelo hitro zbiranje, odkritje in/ali analiziranje (Olofson, Vesset, 2012: 4). Ta definicija obsega strojno in programsko opremo ter storitve, ki integrirajo, organizirajo, upravljajo, analizirajo in prikazujejo podatke. V angleškem jeziku temeljne značilnosti velikega podatkovja na kratko opredeljujejo s »štirimi V-ji«: *volume* (količina), *varietiy* (raznolikost), *velocity* (hitrost) in *value* (vrednost). Prve tri značilnosti (količina, raznolikost in hitrost) so tehnične narave in so odvisne od razvoja kapacitet za shranjevanje in obdelavo podatkov. Četrta značilnost (vrednost) pa označuje potencialno družbeno-ekonomsko vrednost, ki je v bistvu tudi pomemben motivacijski (in investicijski) dejavnik za zbiranje, obdelovanje in uporabo podatkov. S tega vidika tudi OECD poudarja, da je na pojem velikega podatkovja primerno gledati širše kot zgolj z vidika tehničnih aspektov in upoštevati tudi družbeno-ekonomsko dimenzijo tega pojava kot »novega proizvodnega faktorja« (OECD, 2013: 12). Dejstvo je, da so možnosti analiz in novih vpogledov, temelječih na velikem podatkovju, vodile v velike investicije na tem področju, ustvarjajo pa tudi pričakovanja, da bo veliko podatkovje razrešilo različne probleme na številnih področjih (Crawford, Schultz, 2014: 96). Koliko so te investicije upravičene in pričakovanja realna, je odvisno tudi od tega, kako bo na spremenjene družbene razmere dolgoročno reagiralo pravo varstva potrošnikov. *Newman* denimo meni, da bi morala pravna ureditev na tem področju 1) okrepiti nadzor posameznikov nad njihovimi osebnimi podatki, 2) uveljaviti strukturne spremembe na trgu za povečanje konkurenčnosti in 3) neposredno regulirati platforme velikega podatkovja z namenom prepovedi škodljivih praks (Newman, 2014: 19–20). To presega ustaljeno ozko gledanje na posamezno pravno področje oz. disciplino, ampak terja širši sistemski pristop, ki ga pravni regulatorji širom po svetu čedalje težje realizirajo tudi zaradi neenotnih konceptov oz. pogledov na človekove pravice, vključno s pravico do zasebnosti in varstva osebnih podatkov.

3.2 Veliki podatki kot osnova novim poslovnim modelom in gonilo gospodarskega razvoja

Ekonomsko gledano predstavljata v dobi velikega podatkovja dodano vrednost oz. potencialni profitabilni element predvsem dve okoliščini:

- povečanje učinkovitosti zaradi samodejne obdelave podatkov in
- možnost analiziranja ogromne količine podatkov, iz katerih je z ustreznimi algoritmi mogoče razbrati potrošniške vzorce in druge

elemente, ki so pomembni za podejtniško odločanje na najrazličnejših področjih (trženje, ciljno oglaševanje, kreditiranje, sklepanje zavarovanj ipd.).

Ključna konkurenčna prednost torej ni zgolj dostop do podatkov, ampak predvsem možnost njihovega kombiniranja in analiziranja ter reagiranje na podlagi ugotovitev. Predvsem na področju trženja oz. marketinga veliko podatkovje tehnično gledano nudi praktično neomejene možnosti, saj omogoča analizo potrošnikovih navad in izdelavo njegovega profila, na podlagi česar lahko podjetja prilagodijo svojo reklamno kampanjo točno določenemu potrošniku. Kot omenjeno, je trend vztrajnega in hitrega povečevanja dostopnosti računalniške tehnologije širokemu krogu ljudi pogojen s čedalje večjo cenovno dostopnostjo te tehnologije. Ena od posledic tega trenda je transformacija tržnih pristopov, ki v čedalje večji meri favorizirajo računalniško podprt direktni marketing, ki temelji na profiliranju potrošnikov, hkrati pa programi za podatkovno rudarjenje zahtevajo obsežne baze podatkov (Mantelaro, 2014: 648). Kombinacija obdelave velikih podatkov in uporabe mobilnih naprav, ki omogočajo analizo geo-lokacijskih podatkov, dalje omogoča, da podjetje potrošniku relevanten oglas dostavi ob ravno pravem času in tudi na ravno pravem kraju (Navetta, 2013: 15). V bistvu je celoten poslovni model interneta zasnovan pretežno na oglaševanju. Internetna podjetja (denimo Amazon, Google, eBay) so med prvimi pričela uporabljati analitiko velikega podatkovja za identificiranje vzorcev obnašanja potrošnikov, čemur ne prilagajajo le oglaševanja, ampak tudi strukturo spletnih strani. Spletne trgovine razpolagajo ne le s podatki, kaj so posamezniki kupili, ampak tudi s podatki, kaj si ogledujejo, kar jim omogoča, da prodajne strategije sproti prilagajajo potencialnim strankam. Tudi finančne institucije analizirajo vzorce obnašanja potrošnikov in jih ustrezno segmentirajo, preden posamezniku ponudijo prav njemu prilagojene finančne produkte. Določena podjetja preverjajo učinkovitost svojih oglaševalskih kampanj z analizo odziva svojih strank na Facebooku, Twiterju in drugih družbenih omrežjih (Donova, Finn, Wadhva, 2014: 27). T. i. BIA (business inteligence and analyitcs) programski trg naj bi bil vreden 16 milijard ameriških dolarjev in naj bi rastel v povprečju 8 % letno, pri čemer smo trenutno priča prestrukturiranju BIA iz retrospektivne BIA, usmerjene v merjenje in poročanje, k napovedni (proaktivni) BIA, usmerjeni v napovedovanje, predvidevanje in modeliranje. Tržni epicenter velikega podatkovja tako postaja ne toliko vedeti več od konkurence, kaj je potrošnik pravkar kupil, ampak kaj bo naslednji potrošnikov nakup oz. korak (Kemp, 2014: 482).

Poslovni modeli, ki temeljijo na velikem podatkovju, so zaenkrat povezani predvsem s trženjem oz. oglaševanjem. Analize ogromnih količin podatkov, povezanih z dogajanjem na trgu najrazličnejših produktov in storitev v kombinaciji s potrošniškimi profili oglaševalcem omogočajo, da stopnjujejo personaliziranost (in pogosto tudi agresivnost) svojih pristopov in povečujejo

dobičke (razprava o tem, ali je to dolgoročno lahko gonilo vzdržnega in zdravega gospodarskega razvoja, presega namen tega prispevka).

3.3 Veliki podatki kot nova vrsta premoženja in tržna kategorija

Koristni učinki velikega podatkovja naj bi glede na naravo tega fenomena (na srečo) preseg(a)li zgolj področje trženja in oglaševanja. Po oceni OECD smo tako trenutno na poti k družbeno-ekonomskemu modelu, ki ga bodo poganjali podatki (data-driven socio-economic model). V tem modelu so osnovno oz. temeljno premoženje podatki, ki lahko ne le ustvarijo občutno konkurenčno prednost, ampak tudi spodbujajo inovacije, vzdržno rast in razvoj. Na področju gospodarstva izkoriščanje podatkov obljublja dodano vrednost pri številnih poslovnih dejavnostih, od optimizacije proizvodnje do bolj učinkovitega izkoriščanja dela ter boljših odnosov s kupci oz. potrošniki. Koristi, ki naj bi jih (poleg izboljšanja trženja s pomočjo ciljnega oglaševanja in personaliziranih priporočil) po pričakovanjih prineslo izkoriščanje podatkov, so v splošnem naslednje (OECD, 2013: 4):

- povečanje raziskav in razvoja,
- razvoj novih produktov in storitev z uporabo podatkov kot produktov ali kot osnovnega elementa produktov,
- optimiziranje proizvodnje ali dobave in
- razvoj novih organizacijskih in upravljavskih pristopov ali občutno izboljšanje obstoječih praks.

Podatki se tako navajajo kot »novo gorivo« informacijske dobe oz. kot premoženje, ki ga korporacije uporabljalo za preoblikovanje trgov in zvišanje svoje tržne moči in dobička. Če pogledamo na primeru: v 6000 trgovinah podjetja Wal-Mart Stores po svetu se vsak dan opravi približno 267 milijonov transakcij. Da bi izboljšalo svojo konkurenčnost, je to podjetje v sodelovanju s podjetjem Hewlett Packard vzpostavilo podatkovno skladišče (data warehouse) s kapaciteto štirih petabajtov (tj. 4000 trilijonov bajtov), v katerem se shrani čisto vsak zapis o nakupih v njihovih trgovinah. Sofisticiran računalniški program, temelječ na strojnem učenju, je podjetju omogočil, da je izboljšalo učinkovitost svoje cenovne politike in oglaševalskih kampanj. Ogromno podatkovno skladišče s pridom uporabljajo tudi službe podjetja za inovacije in dobavno verigo (Chen, Zhang, 2014: 316).

Platforme velikega podatkovja, ki zbirajo stalno naraščajoče podatke o obnašanju, potrebah in interesih potrošnikov, so sicer v porastu predvsem na internetu (Google, Amazon, Apple, Facebook). Medtem ko ti podatki nedvomno povečujejo dobičke gospodarstva vsaj zaradi novih trženjskih pristopov, pa je vprašanje, ali so tudi potrošniki zaradi tega na boljšem. Kot opozarja *Newman*, povečana izguba nadzora nad svojimi podatki posameznike izpostavlja možnostim novih oblik ekonomskega izkoriščanja. Mnoge »brezplačne« storitve na internetu zbirajo podatke, ki jih posredujejo tretjim osebam, primarno oglaševalcem, algoritmično

profiliranje pa podjetjem omogoča razvrščanje potrošnikov, ki je potencialno lahko tudi diskriminatorno. Profiliranje denimo omogoča prilagajanje cene posameznim potrošnikom na način, da potrošniki za isto blago ali storitve plačajo različno ceno glede na skupino, v katero so razvrščeni oz. profilirani (podjetje prodajno politiko naravna tako, da iz vsake transakcije izvleče maksimum). Ta cenovna diskriminacija je sicer lahko potrošnikom v korist, vendar le, če so jim dostopne informacije o ceni, ki jo za isto blago ali storitve plačujejo drugi, in dana možnost, da pridejo do blaga ali storitve po najnižji možni ceni, kar pa pogosto ni primer. Ekonomski izračuni kažejo, da cenovna diskriminacija na daljši rok pomeni višanje cen za vse potrošnike (Newman, 2014: 12–13, podobno tudi Jerome, 2014: 219–220). Profiliranje in stalno spremljanje posameznika lahko ima zanj tudi druge negativne učnike. V literaturi se navaja primer poslovneža, ki je potem, ko se je vrnil s poročnega potovanja po eksotičnih krajih, ugotovil, da mu je banka znižala limit na kreditni kartici iz 10.800 na 3.800 USD, kar je utemeljila s tem, da je kupoval v trgovinah, katerih lastniki imajo »slabo plačilno zgodovino« (Newman, 2014: 14).

Podatkovne baze posameznih podjetij (ne le spletnih, ampak tudi navadnih, ki podatke zbirajo npr. preko kartic oz. programov zvestobe), so vsekakor velik kapital, saj razkrivajo navade, življenjski slog, status, finančni položaj, lahko pa tudi zdravstveno stanje ipd. potrošnikov. Kot taki, podatki postajajo pomembno tržno blago, pri čemer pravni koncepti, ki razpolaganje s podatki pogojujejo s soglasjem posameznika, v praksi niso pretirano uspešni oz. delujejo le tam, kjer podjetja nimajo interesa nad trženjem podatkov, ampak jih nasprotno hranijo kot svojo poslovno skrivnost (npr. podatke o kupcih, vključenih v programe zvestobe). Uporabniki Facebooka, ki ne preberejo pogojev uporabe (kar ni redek pojav), pogosto sploh ne vedo, da so dovolili, da Facebook spremlja, zbira, hrani in obdeluje njihove internetne aktivnosti, ko so prijavljeni v sistem. Ti podatki se nato prodajajo oglaševalcem, ki posamezniku ob obisku določene spletne strani ponudijo njemu prilagojene oglase (Nelson, Simek, 2013: 25).

3.4 Splošno o stroškovni analizi

Kljub velikemu napredku tehnologije strošek zbiranja in obdelave velike količine podatkov še vedno ni povsem zanemarljiv. Poleg obsežnih kapacitet za hrambo podatkov je običajno visok tudi strošek dela, saj je za ustvarjanje analitičnih algoritmov potreben usposobljen kader z ne le tehničnim, ampak tudi analitičnim znanjem. Pri ogromni količini podatkov, ki so na voljo, je namreč zelo pomembna tudi identifikacija podatkov, ki jih je sploh smotrno zbirati, obdelovati oz. meriti zaradi izboljšanja poslovnih odločitev (Olofson, Vesset, 2012 : 4). V stroškovni analizi je torej treba upoštevati na eni strani strošek strojne in programske opreme za zbiranje, hrambo, analizo in tudi varovanje[4] podatkov ter strošek kadra, ki je za to usposobljen, na drugi strani pa vrednost samih podatkov oz. sporočil ali trendov, ki so razvidni iz teh podatkov.

Kot razvidno zgoraj, je trenutno glede poslovnih modelov, ki temeljijo na velikem podatkovju, največ napisanega in znanega o tistih, ki se uveljavljajo na širšem področju oglaševanja in trženja oz. prilagajanja prodajnih strategij točno določenemu potrošniku. Iz tega je mogoče sklepati, da je stroškovna analiza tovrstne uporabe podatkov pozitivna oz. da se, enostavno rečeno, splača. Kot že rečeno, razprava o tem, ali so na spodbujanju divje potrošnje utemeljeni poslovni modeli lahko gonilo vzdržnega in zdravega gospodarskega razvoja, presega namen tega prispevka; se pa (laično gledano) dolgoročno kot »prava stvar« bolj kaže povečanje vlaganj v raziskave in razvoj inovacij na temelju uporabe podatkov kot produktov ali kot osnovnega elementa produktov. V tem segmentu je stroškovna analiza praviloma bolj zapletena, še posebej v manj dobro stoječih gospodarstvih, kjer so podjetja zainteresirana le za končne produkte, ki jih lahko čim prej prodajo, ne pa za vlaganja v raziskave in razvoj.

Ekonomske prednosti velikega podatkovja niso vezane le na zasebni sektor, ampak se z analizo ustreznih podatkov lahko precej povečata produktivnost in učinkovitost v javnem sektorju. Javni sektor je zelo »podatkovno produktiven«, saj ustvarja velike količine podatkov na najrazličnejših segmentih svojega delovanja, kar pomeni znatne možnosti za vsebinsko najrazličnejše analize. Kar se tiče stroškovne analize, je pri tem treba ustrezno upoštevati, da cilj delovanja institucij javnega sektorja ni ustvarjanje dobička, zato so tudi merske enote drugačne in lahko vključujejo tudi kategorije, kakršne so izboljšana kvaliteta življenja posameznikov, večja varnost ipd.

3.5 Razmerje med pričakovanji in dejanskim stanjem

Kot ugotavlja *Mattioli*, je potencial velikega podatkovja kljub poudarjenim pričakovanjem zaenkrat v veliki meri neizpolnjen. V strokovni in znanstveni literaturi se pretežno navajajo eni in isti primeri, s katerimi se kaže moč velikega podatkovja, vendar so ti primeri dejansko bolj izolirani eksperimenti kot pa dokaz za obširno industrijsko in znanstveno aktivnost. Med bolj pogostimi je študija, ki so jo leta 2010 izvedli raziskovalci na univerzi Stanford skupaj z Microsoftom in v okviru katere so razvili nov način za napovedovanje škodljivega medsebojnega učinkovanja zdravil. Namesto uveljavljenih metod napovedovanja na podlagi študij o kemijskem součinkovanju in človeški fiziologiji je raziskovalna skupina kot vir podatkov vzela internet. Raziskovalci so analizirali podatke iz več milijonov *online* iskanj, izvedenih preko iskalnikov Google, Bing in Yahoo! in s pomočjo statističnih tehnik opazili, da so posamezniki, ki so iskali po imenih dveh zdravil (Paxil in Pravastatin), pogosto v iskalni vnesli tudi iskalni pojem, povezan s hipoglikemijo. Na podlagi te ugotovitve so raziskovalci postavili hipotezo, ki je bila kasneje eksperimentalno potrjena, da imata zdravili Paxil in Pravastatin, če se ju jemlje skupaj, nezaželene stranske učinke (Mattioli, 2014: 540–541). Kot tipičen primer komercialne oblike rabe velikega podatkovja in ponazoritev njegovega (ne vedno pozitivnega) dometa pa se v literaturi pogosto navaja primer trgovske verige Target, ki jo je oče mladoletnega dekleta obtožil nelojalne prakse,

ker je njegovi hčerki pošiljala kupone za opremo za dojenčke in nosečnice. Dejansko je bilo dekle noseče, vendar je trgovina to uspela ugotoviti pred starši. Trgovine si namreč prizadevajo še pred rojstvom otrok identificirati bodoče družine, saj te v povezavi z rojstvom otroka veliko kupujejo, hkrati pa po rojstvu otroka občutno spremenijo potrošniške navade in to »okno priložnosti« želijo trgovci izkoristiti čim prej[5] (Završnik, 2014: 41, tudi Jerome, 2014: 230–231; Joh, 2014: 36–37).

Zakaj torej veliko podatkovje še ni sprožilo velikih inovacij in ekonomske ekspanzije, ki jo konstantno predvidevajo in napovedujejo že nekje od leta 2010? Razlog je delno tehnične, delno pa tudi pravne narave. Večina razprav, ki obravnavajo izjemen potencial velikega podatkovja, izhaja iz predpostavke, da je podatke mogoče enostavno ponovno uporabiti in jih (ponovno) kombinirati zaradi vedno novih analiz. Dejansko stanje pa je drugačno, saj obstajajo številne ovire za enostavno ponovno uporabo podatkov.[6] En sklop teh ovir je tehnične narave in izhaja iz dejstva, da so podatki pogosto zabeleženi in objavljeni v povsem različnih formatih oz. oblikah, zato imajo analitiki oz. raziskovalci težave pri njihovem združevanju iz različnih virov. Ta problem bo sčasoma rešen, saj mednarodne organizacije za standardizacijo že razvijajo modele za standardizirane oblike podatkov, ki bodo omogočili lažje združevanje podatkov za potrebe analitike (Mattioli, 2014: 544–545). Kar se tiče prava, pa je vse odvisno od tega, ali bo uspelo prepoznati dejanska vprašanja in izzive ter reagirati na način, ki bo dolgoročno vzdržen tudi v dobi izjemno hitrega tehnološkega razvoja.

4 Izbrani pravni vidiki velikega podatkovja

4.1 Splošno

Kot že navedeno, internet, veliko podatkovje, digitalno vesolje, kibernetski prostor in druge tehnološke inovacije oz. paradigme, ki bistveno opredeljujejo sodobno družbo in so postali ali postajajo običajen del vsakdanjega življenja, v pravu še nimajo sistematičnega odziva, niti pravega mesta. Definiranih je veliko izzivov in odprtih vprašanj, odgovorov in novih konceptov pa je malo. To je verjetno tudi posledica različnih pogledov na vsebino in razsežnosti človekovih pravic (in v tem okviru še posebej na pravico do zasebnosti in varstva osebnih podatkov) v anglo-saksonskem (predvsem ameriškem) in evropskem kontinentalnem pravnem prostoru. Tudi sicer je pravo tradicionalno počasno v odzivanju na družbene in tehnološke spremembe. Tehnološki napredek, ki smo mu priča v zadnjih nekaj desetletjih, je zgodovinsko gledano razmeroma unikaten, kar se tiče hitrosti, s katero je sodobna informacijsko-telekomunikacijska tehnologija prodrla v vse pore družbe in življenja, in zanesljivo unikaten, kar se tiče količine generiranih podatkov, kar v kombinaciji z dejstvom, da ga poganjajo naravoslovne znanosti, postavlja pravo (pa tudi druge družboslovne znanosti) pred zahtevne izzive. Če se omejimo le na veliko podatkovje, bi bilo s sistemskega vidika bistveno najprej odgovoriti na vprašanje, ali veliko podatkovje ustvarja nove pravne izzive ali pa

gre le za stare oz. že znane pravne probleme v večji preobleki, in ali pojav velikega podatkovja terja zakonski odziv ali pa so potrebni le inovativni pristopi pri interpretaciji oz. razlagi pravnih pravil, nato pa na podlagi ugotovitev iskanje ustreznih modelov in rešitev.

Navedena vprašanja so (sicer parcialno) podrobneje obravnavana v nadaljevanju prispevka. Že na tem mestu pa je treba zapisati, da dokončnih in sistemskih odgovorov na področju pravnih vidikov uporabe velikega podatkovja (predvsem v zasebnem sektorju) ne bo, dokler ne bo (na nadnacionalni ravni in zavezujoče) rešeno temeljno vprašanje, čigavi so podatki, ki jih posameznik generira s svojimi *online* oz. digitalnimi aktivnostmi. Na tem področju se bije boj med gospodarstvom na eni in potrošniki na drugi strani, v katerega so vključeni tudi regulatorji oz. zakonodajalci, ki bodo v končni fazi morali postaviti pravila. *Pentland* v zvezi s tem govori[7] o novi pogodbi o podatkih (new deal on data), v okviru katere se zavzema za jasno opredelitev dejstva, da so podatki, ki jih generira, last posameznika, na katerega se nanašajo, ne glede na to, kdo te podatke zbira, se pravi, da ima posameznik absolutno pravico nad temi podatki enako kot ima absolutno pravico nad svojim telesom in denarjem. V tem konceptu je bistvena transparentnost glede vseh okoliščin v zvezi z zbiranjem podatkov, pri čemer je končni cilj ta, da bi imel dostop do celotne slike, ki jo veliko podatkovje izriše o njem, le posameznik, na katerega se podatki nanašajo (in ne npr. Google, Facebook, kreditne institucije, zavarovalnice ipd.). *Pentland* ob tem še navaja, da raziskave kažejo, da so posamezniki v sistemih, v katerih je zbiranje podatkov regulirano in posledično natančno vedo, kaj se z njihovimi podatki dogaja, hkrati pa sistemu zaupajo, da podatke ustrezno varuje, pripravljeni razkriti več osebnih podatkov kot v sistemih, kjer tega ni.[8]

Penlandova ideja pri internetnih podjetjih in na oglaševanju temelječi industriji seveda ni bila sprejeta,[9] pa tudi na ravni regulatorjev trend ne gre v njeni smeri. Iz uvodnih točk 51 in 57 Predloga Uredbe Evropskega Parlamenta in Sveta o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov (COM/2012/011; v nad.: Predlog Splošne uredbe o varstvu podatkov) denimo izhaja, da bi

> *'morala imeti vsaka oseba pravico dostopa do podatkov, ki so bili zbrani v zvezi z njo, in do enostavnega uveljavljanja te pravice, ter pravico, da bi se seznanila z obdelavo svojih podatkov in preverila njeno zakonitost. Zato bi moral vsak posameznik, na katerega se nanašajo osebni podatki, imeti pravico do seznanitve zlasti o namenih obdelave podatkov, obdobju obdelave, prejemnikih podatkov, logiki podatkov, ki se obdelujejo, in možnih posledicah obdelave, vsaj v primerih, kadar podatki temeljijo na oblikovanju profilov. Ta pravica na drugi strani ne bi smela škodljivo vplivati na pravice in svoboščine drugih, vključno s poslovnimi skrivnostmi ali intelektualno lastnino, in predvsem na avtorske pravice, ki ščitijo programsko opremo. Vendar pa to ne sme povzročiti, da se*

> *posamezniku, na katerega se nanašajo osebni podatki, zavrnejo vse informacije. […]Kadar se njegovi osebni podatki obdelujejo zaradi neposrednega trženja, bi moral imeti posameznik pravico, da taki obdelavi ugovarja brezplačno in na način, ki ga je mogoče enostavno in učinkovito uveljavljati.'*

Kar se tiče profiliranja, ki z velikim podatkovjem dobiva naslutene razsežnosti, pa iz uvodne točke 58 Predloga splošne uredbe o varstvu podatkov izhaja, da bi

> *'vsaka fizična oseba morala imeti pravico, da ni predmet ukrepa, ki temelji na oblikovanju profilov s samodejno obdelavo podatkov. Vendar pa bi moral biti tak ukrep dovoljen, kadar ga izrecno dovoljuje zakon, kadar je sprejet med sklepanjem ali izvrševanjem pogodbe ali kadar je posameznik, na katerega se nanašajo osebni podatki, dal privolitev. V vsakem primeru bi morali za take postopke veljati ustrezni zaščitni ukrepi, vključno s posebnim obveščanjem posameznika, na katerega se nanašajo osebni podatki, in pravico do človeškega posredovanja, veljati pa bi moralo tudi, da se tak ukrep ne sme nanašati na otroka.'*

Podobno Predlog Direktive Evropskega parlamenta in Sveta o varstvu posameznikov pri obdelavi osebnih podatkov, ki jih pristojni organi obdelujejo za namene preprečevanja, preiskovanja, odkrivanja ali pregona kaznivih dejanj ali izvrševanja kazenskih sankcij, in o prostem pretoku takih podatkov (COM/2012/010; v nad. Predlog Direktive o varstvu posameznikov pri obdelavi podatkov) v uvodni točki 27 poudarja, da bi morala

> *'imeti vsaka fizična oseba pravico, da se v zvezi z njo ne sprejme ukrep, ki temelji izključno na samodejni obdelavi podatkov, če ima škodljiv pravni učinek na navedeno osebo, razen če tako določa zakon in če je sprejet z ustreznimi ukrepi za varovanje pravnega interesa posameznika, na katerega se nanašajo osebni podatki.'*

V duhu teh načel sta oblikovana tudi predloga besedil nove uredbe in direktive. Niso torej predvidene revolucionarne rešitve, ki bi zahtevale univerzalno zavezo, da bo tehnologija (in ekonomija) vedno primarno v službi človeštva oz. človečnosti, ampak se iščejo kompromisi.

4.2 Veliko podatkovje: novi pravni izzivi ali stari pravni problemi v večji preobleki

4.2.1 Splošno

Pravni področji, ki sta v dobi velikih podatkov še posebej izpostavljeni, ko gre za sistemska vprašanja, sta pravo intelektualne lastnine in pravo varstva osebnih podatkov in zasebnosti. Veliko podatkovje sproža različna vprašanja tudi na drugih pravnih področjih, saj masovno shranjevanje najrazličnejših podatkov o posamezniku omogoča posege tudi v druge človekove pravice (npr. pravico do

dostojanstva, do enake obravnave ne glede na raso, spol, vero ipd.,), pred izzivom (oz. problemom) se pogosto znajdejo načela o krajevni veljavnosti zakonodaje, z vidika procesnega prava nastajajo vprašanja, povezana s pristojnostjo sodišč oz. drugih organov, v dokaznem pravu se je pojavila nova kategorija t. i. elektronskih dokazov, ki zahtevajo posebno pozornost pri oceni pravne dopustnosti in verodostojnosti ipd.

Veljavni pravni okvir na mednarodni (nadnacionalni) in nacionalni ravni se ne nanaša neposredno na veliko podatkovje, pač pa na podatke kot takšne, večina pravnih vprašanj pa je tako ali drugače povezanih s postopkom obdelave in ponovne uporabe ogromne mase podatkov, ki se generirajo tudi pri čisto vsakdanjih opravilih, in analiziranjem tako obdelanih podatkov. Pri tem je večina veljavnih aktov nastala še v času t. i. tiskane dobe, v kateri so bile fizične in tehnične omejitve znatno drugačne (večje) kot so v digitalni dobi. EU je denimo prepoznala potrebo, da prenovi svoj pravni okvir na področju varstva podatkov. Trenutno veljavni akti EU na tem področju so naslednji:

- Direktiva Evropskega parlamenta in Sveta 95/46/ES z dne 24. oktobra 1995 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov (v nad.: Direktiva 95/46/ES),[10]
- Direktiva 2002/58/ES Evropskega parlamenta in Sveta z dne 12. julija 2002 o obdelavi osebnih podatkov in varstvu zasebnosti na področju elektronskih komunikacij (Direktiva o zasebnosti in elektronskih komunikacijah)[11] in
- Uredba (ES) št. 45/2001 Evropskega parlamenta in Sveta z dne 18. decembra 2000 o varstvu posameznikov pri obdelavi osebnih podatkov v institucijah in organih Skupnosti in o prostem pretoku takih podatkov.[12]

Med načeli, ki jih v zvezi z varstvom osebnih podatkov uveljavlja Direktiva 95/46,[13] se z velikimi podatki še posebej ne ujemata načelo minimizacije podatkov, po katerem se smejo zbirati in obdelovati le tisti osebni podatki, ki so nujni za dosego namena, zaradi katerega se zbirajo oz. obdelujejo, in načelo omejenega namena obdelave podatkov, skladno s katerim smeta zbiranje in obdelava podatkov zasledovati le vnaprej določen, izrecen in legitimen namen. Navedenim vprašanjem je posvečena posebna pozornost v fazi sprejemanja dveh novih aktov, ki bosta posodobila sistem varstva (osebnih) podatkov na ravni EU in ga (do določene mere) tudi uskladila s tehnološkim napredkom. To sta že zgoraj omenjeni:

- Predlog Splošne uredbe o varstvu podatkov,
- Predlog Direktive o varstvu posameznikov pri obdelavi podatkov.

Iz Predloga Splošne uredbe o varstvu podatkov na načelni ravni izhaja, da je obdelava osebnih podatkov namenjena temu, da služi človeku; prispevati pa mora tudi k oblikovanju območja svobode, varnosti in pravice ter gospodarske unije, h gospodarskemu in družbenemu napredku, krepitvi in uskladitvi gospodarstev na

notranjem trgu ter blaginji posameznikov. Bistvenega pomena je, da se oblikuje zaupanje, ki bo digitalnemu gospodarstvu omogočilo razvoj na notranjem trgu. Posamezniki morajo imeti nadzor nad lastnimi osebnimi podatki, pravna in praktična varnost posameznikov, gospodarskih subjektov in javnih organov pa morata biti okrepljeni.

Predlog Direktive o varstvu posameznikov pri obdelavi podatkov pa na načelni ravni izhaja iz dejstva, da sta hiter tehnološki razvoj in globalizacija prinesla nove izzive za varstvo osebnih podatkov, da se je obseg zbiranja in izmenjave podatkov bistveno povečal, tehnologija pa pristojnim organom omogoča, da v doslej največjem obsegu uporabljajo osebne podatke za izvajanje svojih dejavnosti, kar vse skupaj zahteva lajšanje prostega pretoka podatkov med pristojnimi organi v EU ter prenosa v tretje države in mednarodne organizacije, pri čemer je treba zagotoviti visoko raven varstva osebnih podatkov.

Navedena nova pravna akta EU bosta pomemben korak naprej, tudi zaradi prizadevanja, da se na celotnem območju EU ključna vprašanja varstva in obdelave osebnih podatkov uredijo poenoteno, vendar pa predstavljata bolj retroaktivno kot proaktivno odzivanje in posledično ne uvajata inovativnih pravnih konceptov. Iz pregleda pravne literature iz različnih področij, ki se ukvarja z vprašanji velikega podatkovja, izhaja, da bodo potrebni novi koncepti in modeli, če želi pravo konstruktivno sodelovati v tehnološko obarvanem družbenem razvoju. Nekaj novejših razmišljanj v tej smeri je predstavljenih v nadaljevanju, vendar pa konkretnejših sistemskih rešitev z izjemo že omenjenega Pentlandovega predloga »nove pogodbe o podatkih« avtorji (še) ne ponujajo.

4.2.2 (Ne)ustreznost modela vnaprejšnje privolitve v dobi velikega podatkovja

Ko podatki v določenem trenutku niso bili več dostopni le državnim ustanovam za javne namene, ampak tudi subjektom zasebnega prava, ki delujejo po tržnih načelih, se je pravo odzvalo z vzpostavitvijo sistema, v katerem podatki posameznikov ne morejo biti uporabljeni v tržne namene brez njihove privolitve, se pravi z razvojem modela »obveščenosti in soglasja« (notice and consent)[14] (Mantelaro, 2014: 649). Ob razmahu velikega podatkovja je ta model postavljen pod vprašaj. Kot že navedeno, s tehnološkega in ekonomskega vidika ideja velikega podatkovja temelji na tem, da se zberejo ogromne količine podatkov, ki se hranijo za nedoločen čas in uporabljajo za vnaprej nedoločene namene, pri čemer število ponovne uporabe ni omejeno. Uspešnost modela velikega podatkovja torej temelji na načelu maksimizacije podatkov (in njihove dolge/trajne hrambe), saj je bistvo uporabnosti velikega podatkovja prav v analizi velike količine čim bolj raznovrstnih podatkov, ki lahko v korelaciji pripeljejo do novih in nepredvidljivih ponovnih uporab podatkov, ki znatno presegajo prvoten namen zbiranja[15] posamezne vrste podatkov. Informirane privolitve za tovrstno uporabo podatkov realno ni mogoče dati oz. bi bila vsebinsko brezpredmetna. S

tem, ko bi posameznik privolil v vnaprej nedoločeno ali nedefinirano obdelavo svojih podatkov, bi dejansko pristal na vsakršno obdelavo in potencialno torej tudi na to, da se podatki uporabljajo na način, ki mu je v škodo. Četudi bi z inovativnimi pravnimi pristopi morda uspeli razviti kriterije in pogoje, pod katerimi bi se prvotno dano soglasje posameznika za zbiranje in obdelavo njegovih osebnih podatkov lahko širilo na nove (in vnaprej nedoločene) načine uporabe, bi še vedno ostal odprt problem zagotavljanja transparentnosti pri zbiranju podatkov,[16] dostopa do lastnih osebnih podatkov in pravice do popravka oz. izbrisa (Donova, Finn, Wadhva, 2014: 49–50).

Nekaj časa je sicer veljalo prepričanje, da je večino težav, povezanih s (ponovno) uporabo podatkov, mogoče rešiti s procesom anonimizacije podatkov že v času njihove obdelave ali pa kasneje v procesu analize. Pravo varstva osebnih podatkov namreč zajema le podatke o določenem ali določljivem posamezniku, se pravi podatke, na podlagi katerih je mogoče posameznika identificirati. Ostali podatki, vključno z anonimiziranimi podatki, so t. i. neosebni podatki, ki ne uživajo posebnega pravnega varstva po teh predpisih. Vendar pa v dobi velikega podatkovja tudi anonimizacija ni enostavna, saj lahko povezava različnih podatkov, ki so vsak zase sicer anonimizirani, razkrije identiteto posameznika (o tem podrobno Ohm, 2010: 1701–1777). Do takšne neželene posledice, za katero se je ustalil izraz »re-identifikacija«, lahko pride povsem nehote, predvsem kadar se podatki povezujejo z vedno novimi podatki, ki prvotno bodisi še niso obstajali ali pa niso bili predvideni za povezovanje. Problem »re-identifikacije« v osnovi ni povezan le z anonimiziranimi podatki, ampak z vsemi vrstami neosebnih podatkov, katerih narava se lahko sčasoma zaradi novih možnosti povezovanja podatkov spremeni in torej iz neosebnih postanejo osebni.[17] Koncept vnaprejšnje privolitve posameznika v takem primeru realno ni izvedljiv. Prav tako ni prav realen v povezavi z računalništvom v oblaku, ki temelji na nizu oz. verigi dejavnosti, ki se prenašajo med različnimi akterji (tj. nosilci platform in infrastrukture), o čemer povprečni končni uporabnik ne ve ničesar, hkrati pa je v tem primeru težko (ali celo nemogoče) določiti, kdo je odgovorni obdelovalec osebnih podatkov.

Ker je nanizane težave ob obstoječih pravnih konceptih nemogoče (raz)rešiti, so nekateri mnenja, da bi bilo treba (vsaj na področju velikega podatkovja) opustiti sistem privolitve posameznika za obdelavo osebnih podatkov in ga nadomestiti z nadzorom s strani neodvisne javne institucije. Konkretneje *Mantelaro* ugotavlja, da posamezniki niso zmožni ali pa nimajo možnosti podrobneje razumeti procesov obdelave velikega podatkovja in namenov te obdelave, zato tudi ne morejo sprejemati svobodnih in ozaveščenih (informiranih) odločitev v zvezi s tem. Posledično pravna ureditev ne bi smela temeljiti primarno na (samo)odločitvi posameznika, ampak bi bilo potrebno vlogo posameznika zmanjšati, hkrati pa okrepiti vlogo neodvisnih nadzornih institucij za varstvo osebnih podatkov. Drugače od posameznikov imajo te institucije zadostno tehnično znanje, da lahko ocenijo tveganja, ki jih prinaša posamezna obdelava podatkov in lahko uporabijo

tudi pravne vzvode za obvladovanje teh tveganj. Nadalje njihov položaj tem institucijam omogoča, da uravnotežijo različne (kdaj tudi nasprotujoče si) interese različnih subjektov na področju obsežnih projektov zbiranja podatkov in podatkovnega rudarjenja. V luči navedenega *Mantelaro* predlaga sprejem pravil za veliko podatkovje in primere odvisnosti od ponudnika storitev oz. potrošnikove priklenjenosti (lock-in), ki bi temeljila na sistemu poznejšega odstopa (opt-out) in hkratni uzakonitvi stroge vsestranske vnaprejšnje ocene tveganj, ki se ne bi nanašala le na obdelavo podatkov, ampak tudi na družbene učnike in etična vprašanja, povezana z uporabo osebnih podatkov. To oceno tveganj bi morala opraviti tretja oseba, nadzor pa bi opravljali pristojni organi za varstvo podatkov, ki bi lahko zahtevali tudi, da se v proces ocene tveganj vključijo različni zainteresirani subjekti (Mantelaro, 2014: 659).

Podobno kot Pentlandova ideja »nove pogodbe o podatkih«, tudi Mantelarov koncept zaenkrat ni bil sprejet. EU denimo ostaja trdno v sistemu *opt-in*, torej vnaprejšnje privolitve, ki bo temeljno načelo tudi po prenovi aktov EU iz področja varstva osebnih podatkov. Kot izhaja iz uvodne točke 29 Predloga Splošne uredbe o varstvu podatkov, bi '*morala biti privolitev dana izrecno s katero koli ustrezno metodo, ki omogoča prostovoljno dano posebno in informirano izjavo volje posameznika, na katerega se nanašajo osebni podatki, tj. z izjavo ali jasnim pritrdilnim dejanjem takega posameznika, kar zagotavlja, da posamezniki vedo, da dajejo privolitev za obdelavo osebnih podatkov […]. Molk ali nedejavnost zato ne bi smela pomeniti privolitve. Privolitev bi morala zajemati vse dejavnosti obdelave, izvedene v isti namen ali namene. Če je privolitev posameznika, na katerega se nanašajo osebni podatki, dana na podlagi elektronske zahteve, mora biti zahteva jasna in natančna, prav tako pa ne sme biti po nepotrebnem moteča za uporabo storitve, za katero se zagotavlja.*' Vendar pa po drugi strani iz uvodne točke 35 Predloga Direktive o varstvu posameznikov pri obdelavi osebnih podatkov izhaja, da mora imeti posameznik v primeru, če je država članica sprejela zakonodajne ukrepe za delno ali popolno omejitev pravice njegove pravice dostopa do podatkov *'pravico, da od pristojnega nacionalnega nadzornega organa zahteva pregled zakonitosti obdelave. Posameznika, na katerega se nanašajo osebni podatki, bi bilo treba o tej pravici obvestiti. Kadar nadzorni organ izvede dostop v imenu posameznika, na katerega se nanašajo osebni podatki, mora nadzorni organ posameznika, na katerega se nanašajo osebni podatki, obvestiti vsaj o tem, da je opravil vsa potrebna preverjanja, in o rezultatu v zvezi z zakonitostjo zadevne obdelave.'* Predvideni so torej novi instituti, vendar večinoma v okviru že znanih sistemskih rešitev.

Ob rob navedenim razpravam je treba zapisati tudi, da posamezniki praviloma ne razmišljajo preveč o svoji digitalni zasebnosti. S tem, ko uporablja GPS v avtomobilu, objavlja sporočila na Facebooku, kupuje na Amazonu, uporablja aplikacije za pametne telefone, ki zahtevajo podatek o njegovi lokaciji, plačuje s kreditnimi karticami ipd. posameznik ustvarja (velik) podatkovni oblak o tem, kdo je, kje je, kaj dela, kaj mu je všeč in kaj ne (Nelson, Simek, 2013: 25). Posledično

v času, ko dnevno tako ali drugače (in pogosto brez možnosti prave izbire[18]) soglašamo z uporabo svojih podatkov, osrednje vprašanje oz. izziv postaja nadzor nad podatki in omejevanje dostopa do njih (in ne več zbiranje podatkov, ker je to že ušlo izpod možnosti nadzora oz. posamezniki z njim praviloma soglašajo oz. morajo soglašati). To potrjuje tudi odmevna sodba Sodišča EU zoper Google,[19] v kateri sodišče ni odredilo izbrisa oz. uničenja spornih podatkov o posamezniku, ampak je omejilo le interes oz. pravico javnosti do dostopa do teh podatkov, pri čemer je Google zavezan to omejitev zagotavljati le na območju EU (podrobnejšo analizo in kritiko te odločitve gl. v Gerry, Berova, 2014: 472–475). Podobno je v uvodni točki 5 Predloga Splošne uredbe o varstvu podatkov ugotovljeno, da

> *'sta hiter tehnološki razvoj in globalizacija prinesla nove izzive za varstvo osebnih podatkov. Obseg izmenjave in zbiranja podatkov se je bistveno povečal. Tehnologija zasebnim podjetjem in javnim organom omogoča, da osebne podatke uporabljajo za izvajanje svojih dejavnosti v obsegu, kakršnega še ni bilo. Posamezniki vedno bolj dajejo osebne podatke na razpolago tako javno kot globalno. Tehnologija je spremenila gospodarstvo in družbeno življenje ter zahteva nadaljnje lajšanje prostega pretoka podatkov v Uniji ter prenosa v tretje države in mednarodne organizacije, pri čemer je treba zagotoviti visoko raven varstva osebnih podatkov'.*

S tega vidika se denimo kot nerazumna kaže ureditev, po kateri lahko trgovska veriga na podlagi privolitve posameznika, ki jo »plača« z nekaj evri popustov, s posameznikovimi podatki počne prav vse, česar se spomni, s primarnim ciljem od njega izvleči čim več denarja, ne smejo pa se na enem mestu (elektronski kartici) povezati podatki iz uradnih evidenc, ki bi posamezniku omogočili hitro in enostavno uveljavljanje pravic in državi že na srednji rok bistveno znižali stroške, hkrati pa pomembno omejili možnosti goljufij in prevar. Argument, da so možne zlorabe, ni prav prepričljiv – ni samo sistem tisti, ki lahko zlorabi posameznika, ampak lahko tudi posameznik zlorabi sistem, zato je na obeh straneh primerno poskrbeti, da bo možnost zlorab minimalna, primeri, ko bo do njih prišlo, pa ustrezno in dosledno sankcionirani.

4.2.3 »Podatkovna pravica« kot nova pravica intelektualne lastnine

Postopki za obdelavo velikega podatkovja in rezultati obdelave se lahko opredelijo kot poslovna skrivnost. Institut poslovne skrivnosti oz. tajnosti je za te postopke primeren tudi zato, ker jih je težko ali celo nemogoče ponoviti na način, da bi se lahko naknadno rekonstruiral postopek, ki je bil uporabljen. Ali bi na tem področju prišlo v poštev tudi patentno varstvo, je vprašljivo, saj ni povsem jasno, ali so postopki obdelave velikega podatkovja dovolj izvirni, da bi bilo zanje možno zahtevati patentno zaščito (Mattioli, 2014: 556). Precejšnje ovire pri obdelavi velikega podatkovja pa lahko ustvarjajo tudi avtorske pravice, ki obstajajo ne le na računalniških programih za obdelavo podatkov, ampak lahko tudi na posameznem podatku ali pa na celoti podatkovni bazi, seveda pa se vzpostavijo le, če je izpolnjen pogoj izvirnosti.[20] Z vidika Bernske konvencije za

varstvo knjiženih in umetniških del[21] je podatkovne baze treba šteti kot zbirke, Pogodba SOIL o avtorski pravici[22] in Sporazum o trgovinskih vidikih pravic intelektualne lastnine (Sporazum TRIPs[23]) pa zagotavljata izrecno avtorsko zaščito zbirk podatkov. Tudi Evropska unija je že leta 1996 sprejela Direktivo 96/9/EC o pravnem varstvu podatkovnih baz[24] (v nad. Direktiva 96/9/EC). Direktiva bazo podatkov v členu 1 opredeljuje kot zbirko neodvisnih del, podatkov ali drugega gradiva, ki je sistematično in metodično razporejeno in individualno dostopno z elektronskimi in drugimi sredstvi, z avtorsko pravico pa se skladno z direktivo varujejo baze podatkov kot take, ki so zaradi izbora ali razporeditve svoje vsebine avtorjeva lastna intelektualna stvaritev (člen 3 Direktive 96/9/EC). Poleg avtorske pravice Direktiva 96/9/EC podatkovnim bazam zagotavlja tudi t. i. *sui generis* varstvo, in sicer so skladno s členom 7 države članice dolžne zagotoviti pravico izdelovalca baze podatkov, pri katerem je prišlo do kakovostno in/ali količinsko znatne naložbe v pridobivanje, preverjanje ali predstavitev vsebine, da se prepreči neupravičeno jemanje izvlečkov in/ali ponovno uporabo celotne vsebine te baze podatkov ali njenega bistvenega dela, ocenjenega kakovostno in/ali količinsko. Poudariti je treba, da se Direktiva 96/6/EC ne uporablja le za elektronske oz. digitalne podatkovne baze, ampak za vse baze.

Izčrpno analizo institutov poslovne skrivnosti, patenta in avtorske pravice v luči velikega podatkovja je opravil *Mattioli* in pri tem ugotovil, da pravo intelektualne lastnine subjektov, ki ustvarjajo baze velikega podatkovja, ne spodbuja v zadostni meri k razkritju svojih metod in praks, ki je ključno za izrabo potenciala velikega podatkovja. To lahko bistveno oteži razvoj inovativnosti, zato bodo na tem področju potrebne nove politike. *Mattioli* v zvezi s tem predlaga skrbno prilagojeno obliko *sui generis* varstva intelektualne lastnine, ki jo poimenuje »podatkovna pravica« (dataright). Ta institut bi bil na voljo prosilcem, ki bi jasno in v popolnosti razkrili opise svojih podatkovnih zbirk in metod za njihovo oblikovanje ter podatke, ustvarjene s temi metodami. Ta nov pravni institut bi bil opredeljen s tremi značilnostmi, ki so tudi sicer lastne pravicam intelektualne lastnine: 1) predmet pravice, 2) izključna pravica na predmetu in 3) niz pravil, ki zagotavljajo ekskluzivnost (Mattioli, 2014: 578 in 583).

4.2.4 Napovedni modeli v boju zoper kriminaliteto

Najbolj otipljiva oz. vidna uporaba velikega podatkovja pri policijskem delu je trenutno uporaba napovednih modelov (t. i. predictive policing), ki temeljijo na računalniških algoritmih, ki skušajo na podlagi analize (čedalje večjega obsega) podatkov o storjenih kaznivih dejanjih napovedati bodoče gibanje kriminalnih aktivnosti. Pristop sicer v osnovi ni nov, saj policija že dolgo išče načine za identifikacijo vzorcev kriminalnih aktivnosti, ki ji pomagajo pri odločitvi, kam usmeriti svoje vire, vendar pa na velikem podatkovju temelječi napovedni modeli omogočajo, da se za različne napovedi o bodočih kriminalnih aktivnostih izkoristi na tisoče podatkovnih točk. Uporaba statističnih in geo-lokacijskih analiz za napovedovanje možnosti kriminala je v zadnjem času torej nadgrajena z

možnostjo uporabe zbirk velikega podatkovja, kar omogoča večjo učinkovitost analitičnih tehnik (posebej kvantitativnih) za identifikacijo verjetnih tarč za policijsko intervencijo, za preprečevanje kaznivih dejanj ali za preiskovanje storjenih kaznivih dejanj (Perry in ostali, 2013: 1).

Tako denimo policija v ameriškem mestu Santa Cruz uporablja računalniški program, ki predpostavlja, da vzorci kriminalnih aktivnosti sledijo vzorcem, ki so podobni popotresnim sunkom. Računalniški algoritem upošteva petletno statistiko kaznivih dejanj, ki vključuje čas, kraj in vrsto dejanja in na tej podlagi računa možnost pojava novega kaznivega dejanja v posameznih ožje določenih delih mesta. Pred vsako izmeno policisti dobijo informacije, ki identificirajo 15 delov mesta z najvišjo možnostjo, da se tam zgodi kaznivo dejanje, ter priporočilo (ne pa tudi navodila), da so bolj pozorni na ta področja. Ta model je bil testno predstavljen v letu 2011 in od takrat dalje policijska statistika v mestu Santa Cruz beleži opazen padec vlomov v primerjavi z obdobji pred uvedbo napovednega modela. Tudi v New Yorku je policija (v sodelovanju z Microsoftom) razvila program, ki 24 ur na dan zbira in analizira podatke iz najrazličnejših virov, razpršenih po mestu, in sicer iz 3000 nadzornih kamer, več kot 200 avtomatskih bralnikov registrskih tablic, 2000 senzorjev sevanja ter iz policijskih podatkovnih baz. Povezava vseh teh podatkov daje policiji sliko o varnostni situaciji v mestu, ki je prej ni bilo mogoče izrisati, in s tem možnosti, da zgodaj identificira potencialne grožnje. Policija ima tako v realnem času dostop do podatkov, ki razkrivajo povezave med osebami, predmeti in kraji na način, ki bi ga kriminalni analitiki lahko tudi spregledali (Joh, 2014: 45–46 in 48–49).

V okviru napovednih modelov so se v praksi razvile štiri temeljne kategorije napovednih modelov (Perry in ostali, 2013: xiv):

- metodologije za napovedovanje kaznivih dejanj (uporabljajo se za napovedovanje kraja in časa, ko obstaja povečana verjetnost storitve kaznivega dejanja),
- metodologije za napovedovanje storilca (uporabljajo se za identifikacijo oseb, za katere je verjetno, da bodo v bodoče izvršile kaznivo dejanje),
- metodologije za napovedovanje profilov storilcev (uporabljajo se za ustvarjanje profilov, ki se ujemajo s storilci določenih vrst kaznivih dejanj),
- metodologije za napovedovanje žrtev kaznivih dejanj (uporabljajo se za identifikacijo skupin ali posameznikov, za katere je verjetno, da bodo v prihodnosti žrtve kaznivih dejanj).

Ti modeli oz. metodologije primarno temeljijo na podatkih o preteklih kaznivih dejanjih, lahko pa upoštevajo tudi podatke iz drugih virov, npr. podatke o plačilnih dneh, lokaciji trgovin z alkoholnimi pijačami, sezonskih selitvah, potencialnih poteh za pobeg ipd. (Joh, 2014: 40), pa tudi analize socialnih omrežij z računalniškimi algoritmi, ki izračunavajo verjetnost pojava kriminalnih aktivnosti na podlagi vloge, ki jo ima posameznik v določenem družbenem omrežju (več o tem Završnik, 2014: 44–45).

Vendar pa je treba opozoriti, da napovedni modeli nikoli niso povsem objektivni, saj oblikovanje osnovnih elementov napovednega računalniškega programa oz. algoritma nujno temelji na človeškem faktorju in terja določeno mero diskrecije. Poleg tega lahko napovedni modeli sprožijo odločitve za neupravičene preiskave, ker se policija preveč zanaša na verjetnostne informacije. Če denimo program napove možnost izvršitve vloma v določeni ulici, lahko to povzroči, da bodo policisti videli sumljivo ravnanje tudi tam, kjer ga dejansko ni (Joh, 2014: 58–59). Da lahko pride z nedomišljeno uporabo analitičnih orodij do hudih kršitev človekovih pravic, dokazuje denimo primer iz ameriške zvezne države Maryland, v katerem je policija izkoristila svoj dostop do podatkov v informacijskih centrih (fusion centers) za izvedbo nadzora nad borci za človekove pravice, mirovnimi aktivisti in nasprotniki smrtne kazni. Tekom nadzora, ki je trajal 19 mesecev, je bilo 53 aktivistov s strani računalniškega algoritma okarakteriziranih za teroriste, vključno z dvema katoliškima nunama in demokratskim kandidatom za lokalno funkcijo. Informacijski center je podatke o teh klasifikacijah posredoval zveznim oblastem za bolj proti drogam, v policijske podatkovne zbirke in tudi v baze NSA, ne da bi bile nedolžne osebe kakorkoli obveščene o tem, da so uvrščene na seznam, ali da bi jim bila dana možnost izjasnitve (Crawford, Schultz, 2014: 104). Tudi *Završnik* opozarja, da je morda najnevarnejša posledica velikega podatkovja ta, da so rezultati lahko pristrani in napačni. Kljub temu, da računalniški algoritem deluje na podlagi podatkovnih nizov, ki so sami zase neproblematični in so lahko zbrani in shranjeni povsem pravilno (po zakonu ali soglasjem), pa v kombinaciji omogoča »nove neto« podatke, ki so potencialno občutljivi in jih ni mogoče vnaprej predvideti. Pravni položaj takšnih aktivnosti je še vedno nejasen (Završnik, 2014: 47).

4.3 Veliko podatkovje: izziv za pravno prakso ali tudi za zakonodajalca

Odgovor na vprašanje, kako s (počasnim) pravom slediti hitremu širjenju obsega podatkov na globalni ravni, ni enostaven in tudi ne more biti enoznačen. Da obstoječi pravni okvir ni prilagojen obdelavi velikega podatkovja, ni sporno; prav tako pa je na podlagi vsega navedenega zgoraj mogoče zapisati, da odprtih vprašanj ni oz. jih ne bo mogoče (raz)rešiti z uporabo kreativnih metod razlage prava, se pravi z ustrezno smiselno uporabo oz. navezavo (aplikacijo) obstoječih pravnih mehanizmov na nove razmere, pač pa bodo (bolje prej kot slej) potrebne korenite oz. sistemske spremembe tudi v zakonodajnem okviru.

Vendar pa mora tudi praksa nenehno slediti spremembam in se v okviru svojega delovnega področja ustrezno odzivati na nove družbene razmere, ki vključujejo tudi pospešen tehnološki napredek. Tako *Ferguson* denimo v obsežni razpravi (Ferguson, 2015: 327–410) na temo identifikacije (potencialno) sumljivih posameznikov na podlagi velikega podatkovja ugotavlja, da bodo morala sodišča vsebino razumnega suma oz. razlogov za sum, ki dovoljuje ukrepanje (npr. identifikacijo posameznika),[25] določiti na novo oz. jo vsaj na novo pretehtati v luči dejstva, da lahko organi odkrivanja in pregona ob predpostavki širokih pooblastil

za dostopanje do podatkov in njihovo povezovanje lažje in hitreje odkrijejo storilce kaznivih dejanj (k temu pripomorejo tudi novitete, kakršne so programje za prepoznavo obraza, tehnologije za biometrično identifikacijo, mobilne komunikacije ipd.). Dodatno bodo nov premislek na področju dokaznih standardov zahtevale možnosti nepredvidenih oz. nepričakovanih ugotovitev, ki se bodo izrisale na podlagi zbrane velike količine sicer na videz povsem nepomembnih podatkov. Za preprodajo mamil storilci potrebujejo plastične vrečke in tehtnico. Za orožje so potrebni naboji. Za vlom v avto je potrebno orodje. S spremljanjem prodaje tovrstnega blaga lahko policija identificira storilce kaznivih dejanj. Podobno je s pranjem denarja, ki se ga morajo tako ali drugače lotiti kriminalne organizacije. Nenavadni depoziti, nakupi in denarne transakcije lahko policiji omogočijo, da odkrije pranje denarja in tudi storilce tega dejanja. Če se dodajo še geografski oz. lokacijski podatki, se lahko izriše vzorec kriminalne dejavnosti. Poznavanje lokacije, kjer se lahko zgodi kaznivo dejanje, pa omogoča bolj učinkovite preventivne dejavnosti. Nove tehnologije, temelječe na velikem podatkovju, omogočajo boljše slednje nacionalnim in mednarodnim kriminalnim dejavnostim, vključno s trgovino z belim blagom, tihotapstvom drog in kartičnimi goljufijami. Nadalje lahko analiza družbenih omrežij razkrije indice o hudodelskem združevanju, izvajanju prostitucije ali kibernetski kriminaliteti (Ferguson, 2015: 396–397) ipd. Praksa mora torej stremeti k temu, da proaktivno izkoristi nove možnosti, ki jih ponuja tehnologija, po potrebi tudi s kreativno pravno interpretacijo obstoječih zakonskih institutov.

5 Zaključek

V Poročilu o pravnih, ekonomskih, družbenih, etičnih in političnih izzivih velikega podatkovja avtorji ocenjujejo, da pravna ureditev ustvarja negotov izid glede konkurenčnosti gospodarstva, saj pravno okolje ostaja kompleksno in podpira pretirano zaščitniške ureditve. Posledično menijo, da obstaja jasna potreba po pravni reformi, s katero bi se na eni strani omogočil izkoristek vseh potencialov velikega podatkovja, po drugi strani pa ustrezno zavarovale vrednote, ki jih ogroža tehnični napredek (Donova, Finn, Wadhva, 2014: 5).

Kljub temu, da je tradicionalni pravni okvir za varstvo osebnih podatkov v krizi, saj novo tehnološko in ekonomsko okolje (tržna koncentracija, družbene in tehnološke spremembe) maje temelje tega okvira,[26] instituti prava varstva osebnih podatkov v splošnem (a s čedalje več izjemami) še preprečujejo, da bi ponudniki blaga in storitev oz. oglaševalci potrošnike časovno in krajevno neomejeno nadlegovali s ponudbami. Če pravo ne bi določalo nobenih pravil in bi lahko npr. spletna podjetja in trgovci brez omejitev uporabljali najnovejše tehnološke dosežke, bi posameznik vsakič, ko bi vstopil v shrambo ali odprl hladilnik, prejel elektronsko sporočilo ali SMS z ugodno ponudbo hrane in pijače, pa zraven še ustrezen kuharski, nutricionističen ali podoben nasvet ipd. – tega pa si večina ljudi verjetno ne želi.

V prispevku so z namenom podpreti tezo o potrebni koreniti pravni reformi parcialno obravnavani nekateri izbrani ekonomski in pravni vidiki velikega podatkovja, seveda pa veliko podatkovje nima pomembnih implikacij le na področju prava in ekonomije, ampak ima tudi pomembne družbene, etične, politične in morda še kakšne učinke, ki jih ni mogoče obravnavati povsem ločeno oz. izolirano, saj gre za kompleksen pojav z učinki na družbo kot celoto. Nova spoznanja in vzorci, ki jih na različnih področjih lahko izrišejo analize ogromnega števila podatkov, lahko vplivajo na večjo produktivnost (predvsem v zasebnem sektorju) in večjo učinkovitost (v zasebnem in javnem sektorju), a hkrati odpirajo številna (predvsem pravna, pa tudi etična) vprašanja. Vsekakor bo za uspešne projekte, temelječe na velikem podatkovju in s tem povezanih tehnologijah, potreben trden (in odločen) analitični pravni okvir, temelječ na razumevanju pravic in dolžnosti, ki izhajajo iz velikega podatkovja, in usmerjen v obvladovanje tveganj in razvoj strukturiranega pristopa k pravno skladnemu in programsko podprtemu procesu zbiranja in obdelave masovnih podatkov (Kemp, 2014: 491).

Sklepno je mogoče zapisati, da med pravom in velikim podatkovjem (in na njem temelječo ekonomijo) zaenkrat še obstaja vzajemen odnos: veliko podatkovje postavlja pravo pred nove izzive, pravna ureditev pa po drugi strani nastavlja ovire masovnemu zbiranju, obdelavi in uporabi tako zbranih podatkov. Kot na večini področij, je tudi tukaj smotrno stremeti k uravnoteženosti, a z jasnim končnim ciljem: tako kot je treba zamejiti težnje, da bi si posamezniki prisvojili dele pravega, fizičnega vesolja (npr. lune in sosednjih planetov), je treba preprečiti tudi, da bo izkoriščanje digitalnega vesolja omogočeno le peščici izbranih subjektov, ki bodo obvladovali tehnologijo (ali sčasoma ona njih). Izziv je torej velik, a tudi težav ne manjka. Ena bistvenih je čas. Tehnologija namreč napreduje znatno hitreje kot pravo. Kot navedeno uvodoma, je nujen sistemski pristop na globalni ravni, če želi biti pravo dolgoročno upošteven dejavnik oz. razumna protiutež tehnološkemu razvoju. Zametkov tovrstnega pristopa še ni opaziti – vprašanje pa je, ali bo pravo še zmožno obvladati tehnologijo, ko bo (končno) dojelo, za kaj gre.

Opombe / Notes

[1] Leta 1965 je *Gordon Moore*, soustanovitelj podjetja Intel, napovedal, da se bo število tranzistorjev, ki jih je mogoče spraviti na en računalniški čip, vsaki dve leti podvojilo. Ta trditev, ki je sčasoma postala znana kot Moorov zakon, še vedno velja, saj se meje tehnologije z novimi materiali in pomanjševanjem tranzistorjev na nanometrsko merilo neprestano premikajo. Moorov zakon je aktualen tudi na področju stroškov, in sicer strošek računalniških izračunov zaradi zmogljivejših računalnikov eksponentno pada (računalniški izračun, ki danes stane 1 dolar, bi pred 50-imi leti stal 10 milijard dolarjev). Danes izdelava tranzistorja za računalniški čip stane manj, kot je strošek tiska ene same črke v časopisu. In če bi avtomobilska tehnologija sledila napredku računalniških procesorjev, bi razdaljo med San Franciscom in New Yorkom (4140 km) z avtomobilom lahko prepotovali v 13 sekundah (povzeto po Levy, 2012: 212–213).

[2] Upoštevati je sicer treba tudi dejstvo, da je dostop do sodobne informacijsko-komunikacijske tehnologije na globalni ravni neenak (v angleškem jeziku ta pojav opredeljuje besedna zveza *digital devide*).

[3] Med temeljnimi je odprtokodni program oz. programska knjižnica Apache™ Hadoop®, ki zagotavlja okvir za obdelavo velikih količin podatkov iz različnih virov oz. računalniških grozdov na podlagi preprostih programskih modelov (več o tem gl. Dai in ostali, 2014: 92–110).

[4] Zelo pomemben vidik, s katerim se ta prispevek sicer ne ukvarja, je vprašanje varnosti podatkov oz. občutljivosti in ranljivosti podatkovnih baz. Ob odsotnosti ustreznega varstva teh podatkovnih baz, ki v veliki meri vsebujejo tudi osebne podatke, se lahko zgodi, da pride do zlorabe teh podatkov v obliki njihove neupravičene prisvojitve ali odtujitve, posega v vsebino ipd.

[5] V povezavi s tem je zanimiv primer *Janet Vertesi*, ki se je odločila, da bo svojo nosečnost skušala prikriti »velikemu bratu« oz. oglaševalcem. Striktno je pazila, da nosečnosti ni omenjala na spletu, poskrbela je tudi, da tega niso omenjali njena družina in prijatelji. Za brskanje o zadevah, povezanih z novorojenčki, je uporabljala anonimni iskalnik Tor. V trgovinah je vse stvari za otroka plačevala z gotovino in brez predložitve kartic zvestobe. Na ta način se je sicer ubranila oglaševalcev, vendar pa se je znašla na listi potencialih storilcev kaznivih dejanj, saj je bilo izključno gotovinsko poslovanje nje in njenega soproga s strani trgovcev naznanjeno policiji (Jerome, 2014: 231).

[6] *Mattioli* denimo ugotavlja, da se strokovnjaki s področja računalništva in informatike pritožujejo, da so podatki, ki vstopajo v sheme velikega podatkovja, pogosto nezadostno dokumentirani in dostopni (razkriti). Nerazkritje izvora podatkov ovira ponovno uporabo podatkov, kar lahko prepreči inovativne aplikacije velikega podatkovja (Mattioli, 2014: 536).

[7] Vir: With Big Data Comes Big Responsibility, An interview with MIT Media Lab's Alex "Sandy" Pentland, Harvard Business Review, November 2014, str. 101–104. Intervju je dostopen tudi na spletni povezavi [https://hbr.org/2014/11/with-big-data-comes-big-responsibility] (obiskano: 8. 2. 2015).

[8] Ibidem.

[9] Največja spletna podjetja (med drugim Google in Cisco) si denimo prizadevajo tudi, da bi bila izvzeta iz nove regulative, ki jo na področju kibernetske varnosti in varnosti omrežij pripravlja Evropska unija. [http://www.reuters.com/article/2014/12/09/us-eu-cybersecurity-idUSKBN0JN26F20141209] (obiskano: 19. 2. 2015).

[10] UL L 281, 23. 11. 1995, stran 355–374.

[11] UL L 201, 31. 7. 2002, stran 37–47.

[12] UL L 8, 12. 1. 2001, stran 1–22.

[13] Osrednja načela varstva osebnih podatkov, ki jih uzakonja Direktiva 95/46, so naslednja (povzeto po Donova, Finn, Wadhva, 2014: 48):

obvezna pravna podlaga za zbiranje in obdelavo osebnih podatkov,
vnaprej določen namen zbiranja in obdelave osebnih podatkov,
načelo minimizacije podatkov,
načelo točnosti in ažurnosti podatkov,
transparentnost obdelave podatkov v razmerju do osebe, na katero se podatki nanašajo,
pravica osebe, na katero se nanašajo podatki, do dostopa do podatkov, do popravka ali izbrisa,
varnost podatkov.

[14] Slovenski Zakon o varstvu osebnih podatkov (Uradni list RS, št. 94/07, ZVOP-1-UPB) temelji na t. i. *opt-in* modelu oz. modelu vnaprejšnje privolitve posameznikov za obdelavo

osebnih podatkov, ki velja, kadar podlaga za obdelavo osebnih podatkov ni določena v zakonu (8. člen ZVOP-1).

[15] Kar se tiče namena zbiranja oz. obdelave podatkov, OECD denimo opozarja, da se obstoječe pravne ureditve praviloma osredotočajo predvsem na to, da mora biti namen obdelave podatkov določen in transparenten pred njihovim zbiranjem in obdelavo, ne vsebujejo pa pravil o tem, za kakšne vrste namenov se smejo (ali ne smejo) uporabljati osebni podatki. Posledično je povsem odprto vprašanje, kje je denimo meja med izboljšanjem odnosa s strankami na eni in nepošteno manipulacijo potrošnikov na drugi strani, ali denimo vprašanje, kdaj optimizacija tveganj postane nepoštena diskriminacija (OECD, 2013: 24).

[16] Zaradi zagotavljanja zanesljivosti rezultatov analiz velikega podatkovja je nujna transparentnost ne le glede virov podatkov, ampak tudi glede metode, po kateri so bili podatki zbrani ter glede morebitnih sprememb podatkov (Mattioli, 2014: 545).

[17] Problem »re-identifikacije« je obravnavala tudi Delovna skupina 29 (WP 29) in v zvezi s tem opozorila predvsem na tri tveganja (WP 29 Opinion 05/2014: 8–10):

neustrezno enačenje pseudonimiziranih podatkov z anonimiziranimi podatki,

miselnost, da ustrezno anonimizirani podatki posamezniku ne zagotavljajo več nobenih varovalk (v zvezi s tem WP 29 opozarja, da se evropska ureditev varstva osebnih podatkov v določenem delu nanaša na vse podatke, tudi t. i. neosebne podatke),

četudi za anonimzirane podatke ne velja evropska ureditev o varstvu osebnih podatkov, pa lahko predvsem v povezavi s profiliranjem in sprejemanjem odločitev, ki (četudi posredno) vplivajo na posameznike, tudi ti podatki sovpadejo s sfero posameznikovega osebnega življenja, ki je varovana z 8. členom Evropske konvencije o človekovih pravicah in 7. členom EU Listine o temeljnih človekovih pravicah.

[18] Argument, da se lahko posameznik pač odloči, da ne bo aktiven v digitalnem svetu, če ga zbiranje podatkov o njegovih aktivnostih oz. vsesplošen nadzor moti (in torej deluje po načelu 'vzemi ali pusti'), danes ni več realen, saj se vse več vsakodnevnih, a nujnih opravil seli v digitalno oz. virtualno okolje, zato ima posameznik precej omejene možnosti odločanja, če želi razmeroma normalno življenje (več o tem Pasquale, Citron, 2014: 1413).

[19] Zadeva C-131/12 z dne 13. maja 2014, *Google Spain SL in Google Inc. proti Agencia Española de Protección de Datos (AEPD) in Mario Costeja González*, še neobjavljena.

[20] Iz razloga, ker ne izpolnjuje pogoja izvirnosti, je Vrhovno sodišče ZDA denimo odločilo, da avtorska pravica ne obstaja na abecednem telefonskem imeniku. Gl. zadevo Feist Pubs., Inc. v. Rural Tel. Svc. Co., Inc., 499 U.S. 340 (1991) [https://supreme.justia.com/cases/federal/us/499/340/] (obiskano: 24. 1. 2015).

[21] Uradni list SFRJ-MP, št. 14/1975 in 4/1986, Uradni list RS št. 24/1992, Uradni list RS-MP, št. 9/1992 in 3/2007.

[22] Uradni list RS-MP, št. 25/1999.

[23] Uradni list RS-MP, št. 10/1995.

[24] UL L 77, 27. 3. 1996, stran 20–28.

[25] Organi odkrivanja in pregona kaznivih dejanj za pričetek svojih aktivnosti potrebujejo določeno stopnjo suma, ki služi kot meja in hkrati kot zaščita posameznika pred nerazumnim policijskim nadlegovanjem. Dokazni standard za pričetek predkazenskega postopka v Sloveniji so razlogi za sum (148. člen Zakona o kazenskem postopku, Uradni list RS, št. 63/1994 in nasl., ZKP), v ZDA glede tega denimo poznajo standard razumnega suma (reasonable suspicion), ki se presoja po kriterijih iz IV. amandmaja k Ustavi ZDA.

[26] Ta okvir temelji predvsem na načelu vnaprejšnje določenosti namena zbiranja podatkov, načelu minimizacije podatkov ter modelu obveščenosti in informirane privolitve posameznika, na katerega se podatki nanašajo (Mantelaro, 2014: 651).

Literatura / References

Chen, P. & Zhang C. (2014) Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, *Information Sciences*, 275, pp. 314–347.

Crawford, K. & Schultz, J. (2014) Big Data and Due Process: toward a framework to redress predictive privacy harms, *Boston College Law Review*, 55(1), pp. 93–128.

Dai, J., Huang, J., Huang, S. & Liu, Y, Sun, Y. (2012) The Hadoop stack: new paradigm for big data storage and processing, *Intel® Technology Journal*, 16(4), pp. 92–110.

Donova, A., Finn, R. & Kush W. (eda.) (2014) Report on legal, economic, social, ethical and political issues, BYTE, available at: http://byte-project.eu/wp-content/uploads/2014/10/BYTE-D2.1_Final_Compressed.pdf (January 22, 2015).

Ferguson, A. F. (2015) Big Data and Predictive Reasonable Suspicion, University of *Pennsylvania Law Review*, 163(2), pp. 327–410.

Gerry, F. & Berova, N. (2014) The rule of law online: Treating data like the sale of goods: Lessons for the internet from OECD and CISG and sacking Google as the regulator, *Computer Law & Security Review*, 30(5), pp. 465–481.

Jerome, J. (2014) Big Data: catalyst for a privacy conversation, *Indiana Law Review*, 48 (1), pp. 213–242.

Joh, E. E. (2014) Policing by Numbers: Big Data and the Fourth Amendment, *Washington Law Review*, 89(1), pp. 35–68.

Kemp, R. (2014) Legal aspects of managing Big Data, *Computer Law & Security Review*, 30(5), pp. 482–491.

Levy, J. (2012) *Čebela v katedrali in še 99 zanimivih primerjav iz sveta znanosti* (Ljubljana: Tehniška založba Slovenije).

Mantelaro, A. (2014) The future of consumer data protection in the E.U.: Re-thinking the "notice and consent" paradigm in the new era of predictive analytics, *Computer Law & Security Review*, 30(6), pp. 643–660.

Mattioli, M. (2014) Disclosing Big Data, *Minnesota Law Review*, 99(2), pp. 535–583.

Murphy, M., Barton (2014) From a Sea of Data to Actionable Insights: Big Data and What It Means for Lawyers, *Intellectual Property & Technology Law Journal*, 26(3), pp. 8–16

Navetta, D. (2013) Legal Implications of Big Data, a Primer, *ISSA Journal*, 11(3), pp. 14–19.

Nelson, S.D., Simek, J. W. (2013) BIG DATA: Big Pain or Big Gain for Lawyers?, *Law Practice: The Business of Practicing Law*. Jul/Aug 2013, 39(4), pp. 24–27.

Newman, N. (2014) How Big Data Enables Economic Harm to Consumers, Especially to Low- Income and Other Vulnerable Sectors of the Population, *Journal of Internet Law*, Dec. 2014, pp. 11–23.

OECD (2013) Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by 'Big Data,' *OECD Digital Economy Papers*, No. 222, OECD Publishing, available at: http://dx.doi.org/10.1787/5k47zw3fcp43-en (January 23, 2015).

Ohm, P. (2010) Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, *UCLA Law Review*, 5 (6), pp. 1701–1777.

Olofson, C. W. & Vesset, D. (2012) White Paper – Big Data: Trends, Strategies and SAP Technology, IDC & SAP, available at: http://www.sap.com/bin/sapcom/de_at/downloadasset.2012-09-sep-26-13.idc-report--big-data-trends-strategies-and-sap-technology-pdf.html (January 23, 2015).

Pasquale, F. & Citron D.K. (2014) Promoting innovation while preventing discrimination: policy goals for the scored society, *Washington Law Reveiw*, 89(4) pp. 1413–1424.

Perry, W. L., McInnis, B., Carter C. P., Smith S. C. & Hollywood, J. S. (2013) *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, RAND Coroporation.

Završnik, A. (2014) Priložnosti in pasti odkrivanja prihodnjega zločina z algoritmi, *Zbornik 7. konference kazenskega prava in kriminologije*, pp. 37–47.